

Robust hamiltonicity of random directed graphs

Asaf Ferber Rajko Nenadov Andreas Noever

asaf.ferber@inf.ethz.ch rnenadov@inf.ethz.ch anoever@inf.ethz.ch

Ueli Peter Nemanja Škorić

upeter@inf.ethz.ch nskoric@inf.ethz.ch

Institute of Theoretical Computer Science
ETH Zurich, 8092 Zurich, Switzerland

Abstract

In his seminal paper from 1952 Dirac showed that the complete graph on $n \geq 3$ vertices remains Hamiltonian even if we allow an adversary to remove $\lfloor n/2 \rfloor$ edges touching each vertex. In 1960 Ghouila-Houri obtained an analogue statement for digraphs by showing that every directed graph on $n \geq 3$ vertices with minimum in- and out-degree at least $n/2$ contains a directed Hamilton cycle. Both statements quantify the robustness of complete graphs (digraphs) with respect to the property of containing a Hamilton cycle.

A natural way to generalize such results to arbitrary graphs (digraphs) is using the notion of *local resilience*. The local resilience of a graph (digraph) G with respect to a property \mathcal{P} is the maximum number r such that G has the property \mathcal{P} even if we allow an adversary to remove an r -fraction of (in- and out-going) edges touching each vertex. The theorems of Dirac and Ghouila-Houri state that the local resilience of the complete graph and digraph with respect to Hamiltonicity is $1/2$. Recently, these statements have been generalized to random settings. Lee and Sudakov (2012) proved that the local resilience of a random graph with edge probability $p = \omega(\log n/n)$ with respect to Hamiltonicity is $1/2 \pm o(1)$. For random directed graphs, Hefetz, Steger and Sudakov (2014+) proved an analogue statement, but only for edge probability $p = \omega(\log n/\sqrt{n})$. In this paper we significantly improve their result to $p = \omega(\log^8 n/n)$, which is optimal up to the polylogarithmic factor.

1 Introduction

A *Hamilton cycle* in a graph or a directed graph is a cycle that passes through all the vertices of the graph exactly once, and a graph is *Hamiltonian* if it contains a Hamilton cycle. Hamiltonicity is one of the central notions in graph theory, and has been intensively studied by numerous researchers. It is well known that the problem of whether a given graph contains a Hamilton cycle is \mathcal{NP} -complete. In fact, Hamiltonicity was one of Karp's 21 \mathcal{NP} -complete problems [12].

Since one can not hope for a general classification of Hamiltonian graphs, as a consequence of Karp's result, there is a large interest in deriving properties that are sufficient for Hamiltonicity. A classic result by Dirac from 1952 [7] states that every graph on $n \geq 3$ vertices with minimum degree at least $n/2$ is Hamiltonian. This result is tight as the complete bipartite graph with parts of sizes that differ by one, $K_{m,m+1}$, is not Hamiltonian. Note that this theorem answers the following question: Starting with the complete graph on n vertices K_n , what is the maximal integer Δ such that for any subgraph H of K_n with maximum degree Δ , the graph $K_n - H$ obtained by deleting the edges of H from K_n is Hamiltonian? This question not only asks for a sufficient condition for a graph to be Hamiltonian, it also asks for a quantification for the “local robustness” of the complete graph with respect to Hamiltonicity.

A natural generalization of this question is to replace the complete graph with some arbitrary base graph. Recently, questions of this type have drawn a lot of attention under the notion of *resilience*.

Roughly speaking, given a monotone increasing graph property \mathcal{P} and a graph or a digraph G which satisfies \mathcal{P} , the *resilience of G with respect to \mathcal{P}* measures how much one must change G , in order to destroy \mathcal{P} . Since one can destroy many natural properties by small changes (for example, by isolating a vertex), it is natural to limit the number of edges touching any vertex that one is allowed to delete. This leads to the following definition of *local resilience*.

Definition 1.1 (Local resilience). *Let \mathcal{P} be a monotone increasing graph property. For a graph G , the local resilience is*

$$r(G, \mathcal{P}) := \min\{r : \exists H \subseteq G \text{ such that} \\ \forall v \in V(G) \, d_H(v) \leq r \cdot d_G(v) \text{ and} \\ G - H \text{ does not have } \mathcal{P}\},$$

while for a digraph G it is defined as

$$r(G, \mathcal{P}) := \min\{r : \exists H \subseteq G \text{ such that } \forall v \in V(G) \\ d_H^+(v) \leq r \cdot d_G^+(v) \text{ and } d_H^-(v) \leq r \cdot d_G^-(v) \\ \text{and } G - H \text{ does not have } \mathcal{P}\}.$$

Sudakov and Vu initiated the systematic study of resilience of random and pseudorandom graphs in [18], and since then this field has attracted substantial research interest (see e.g. [2, 3, 4, 6, 8, 13, 14]).

Let us denote with \mathcal{HAM} the graph property of containing a Hamiltonian cycle (directed, in case of digraphs). Lee and Sudakov [14] proved that for $p = \omega(\log n/n)$, a typical $G \sim G(n, p)$ satisfies $r(G, \mathcal{HAM}) \in (1/2 \pm o(1))$. Note that this result is asymptotically optimal not only with respect to the constant $1/2$ but also with respect to the probability p , since it is well known that a typical graph $G \sim G(n, p)$ is not Hamiltonian for $p = o(\log n/n)$ (see [5]).

For a positive integer n and $0 \leq p = p(n) \leq 1$, let $D(n, p)$ denote the binomial probability space of random digraphs on the set of vertices $[n] = \{1, \dots, n\}$. That is, an element $D \sim D(n, p)$ is generated by including each of the $n(n-1)$ possible ordered pairs of $[n]$ with probability p , independently at random. For this model, Frieze [9] showed that a typical digraph $D \sim D(n, p)$ is Hamiltonian for $p \geq (\log n + \omega(1))/n$. Therefore, it is natural to ask for an analogue to the result of Lee and Sudakov [14] for random digraphs with these densities.

As a first step towards this goal, Hefetz, Steger and Sudakov proved in [10] the following theorem, which is asymptotically optimal with respect to the resilience but far from optimal with respect to the edge probability.

Theorem 1.2 ([10]). *Let $\beta > 0$, let n be a sufficiently large integer and let $p = \omega(\log n / \sqrt{n})$. Then w.h.p. a digraph $G \sim D(n, p)$ satisfies $r(G, \mathcal{HAM}) \in (1/2 \pm \beta)$.*

In the proof of Theorem 1.2 Hefetz, Steger and Sudakov extensively used the Regularity Lemma and the fact that for $p = \omega(\log n / \sqrt{n})$, a typical digraph $G \sim D(n, p)$ contains “many” transitive triangles touching each vertex. Therefore, generalizing it to smaller values of p would at least require to replace triangles with some sparser gadgets.

In general, problems related to Hamilton cycles in digraphs are known to be much harder than their counterparts in the undirected setting, mainly since the Posá rotation-extension technique (see [16]) is, in its simplest form, not applicable to directed graphs.

In this paper we use the absorbing method, initiated by Rödl, Ruciński and Szemerédi [17], combined with a very nice and recent embedding argument of Montgomery [15] to prove the following theorem, which is optimal up to polylogarithmic factors.

Theorem 1.3. *Let $\beta > 0$, let n be a sufficiently large integer and let $p = \omega(\log^8 n / n)$. Then w.h.p. a digraph $G \sim D(n, p)$ satisfies $r(G, \mathcal{HAM}) \in (1/2 \pm \beta)$.*

We want to remark that our proof can easily be turned into a simple and efficient randomized algorithm which finds a Hamilton cycle in a digraph with certain pseudorandom properties.

The paper is organized as follows. In Section 2 we present auxiliary lemmas which are used throughout the paper. In Section 3 we give the definition of (n, α, p) -pseudorandom digraphs and state our main result (Theorem 3.2) concerning the Hamiltonicity of such digraphs. We then show how it implies Theorem 1.3 and furthermore derive the proof of Theorem 3.2 using Connecting and Absorbing lemmas. In Section 4 we then give a proof of Connecting Lemma, and finally in Section 5 we furthermore use it to prove Absorbing Lemma.

1.1 Notation and definitions

For an integer n , let $[n] = \{1, \dots, n\}$ and for $a, b, c \in \mathbb{R}$, let $(a \pm b)c = ((a - b)c, (a + b)c)$.

Our graph theoretic notation is standard and follows that of [19]. In particular we use the following: Given a digraph D we denote by $V(D)$ and $E(D)$ the sets of vertices and arcs of D , respectively, and denote $v(D) = |V(D)|$ and $e(D) = |E(D)|$. For a subset $S \subseteq V(D)$, we denote with $D[S]$ the subgraph of D induced by S . For two (not necessarily disjoint) subsets $X, Y \subseteq V(D)$, set $E_D(X, Y) := \{(x, y) \in E(D) : x \in X \text{ and } y \in Y\}$ and let $e_D(X, Y) = |E_D(X, Y)|$. Furthermore, let $N_D^+(X, Y) = \{y \in Y : x \in X \text{ and } (x, y) \in E(D)\}$ denote the set of all out-neighbors of X in Y and let $N_D^-(X, Y) = \{y \in Y : x \in X \text{ and } (y, x) \in E(D)\}$ denote the set of all in-neighbors of X in Y . Given a vertex $x \in V(D)$ and $\tau \in \{+, -\}$, we abbreviate $N_D^\tau(\{x\}, Y)$ to $N_D^\tau(x, Y)$ and define $d_D^\tau(x, Y) = |N_D^\tau(x, Y)|$ and $d_D^\pm(x, Y) = \min\{d_D^+(x, Y), d_D^-(x, Y)\}$. We omit the subscript D whenever there is no risk of confusion.

For $\tau \in \{+, -\}$ we denote with $\bar{\tau}$ the opposite sign. Furthermore, for $\sigma \in \{+, -\}^\ell$ and $i \in [\ell]$, let $\sigma(i)$ denote i -th member of the ℓ -tuple σ , let $\sigma^i = (\sigma(1), \dots, \sigma(i))$ and let $\bar{\sigma}$ denote $(\bar{\sigma}(\ell), \dots, \bar{\sigma}(1))$. We call a sequence of vertices $P = v_1, \dots, v_{\ell+1}$ a σ -walk if all the vertices are different, except that v_0 and $v_{\ell+1}$ can be the same vertex, and if $v_{i+1} \in N^{\sigma(i)}(v_i)$ for all $1 \leq i \leq \ell$. Moreover, we say that P connects v_1 to $v_{\ell+1}$ and call v_1 and $v_{\ell+1}$ its left and right endpoint, respectively. The σ -walk P is additionally called an $v_1 v_{\ell+1}$ -path if $v_1 \neq v_{\ell+1}$ and $\sigma(i) = +$, for all $1 \leq i \leq \ell$.

2 Tools and preliminaries

In this section we introduce tools used in the proofs of our results.

2.1 Probabilistic tools

We need to employ standard bounds on large deviations of random variables. We mostly use the following well-known bound on the lower and the upper tails of the Binomial distribution due to Chernoff (see [1], [11]).

Lemma 2.1. *Let $X \sim \text{Bin}(n, p)$ and let $\mu = \mathbb{E}(X)$. Then*

- $\Pr[X < (1 - a)\mu] < e^{-a^2\mu/2}$ for every $a > 0$;
- $\Pr[X > (1 + a)\mu] < e^{-a^2\mu/3}$ for every $0 < a < 3/2$.

Remark 2.2. *The conclusions of Lemma 2.1 remain the same when X has the hypergeometric distribution (see [11], Theorem 2.10).*

The following is a trivial yet useful bound.

Lemma 2.3. *Let $X \sim \text{Bin}(n, p)$ and $k \in \mathbb{N}$. Then the following holds:*

$$\Pr(X \geq k) \leq \left(\frac{enp}{k}\right)^k.$$

Proof. $\Pr(X \geq k) \leq \binom{n}{k} p^k \leq \left(\frac{enp}{k}\right)^k.$ □

2.2 Graph Partitioning

The next lemma states that one can partition a digraph into subsets which, proportionally, inherit the lower bound on the in- and out-degree.

Lemma 2.4. *Let $c, \varepsilon > 0$ be constants, n sufficiently large integer and $0 < p := p(n) < 1$. Suppose that:*

- (i) D is a digraph on n vertices,
- (ii) $U \subseteq V(D)$,
- (iii) $k, s_1, \dots, s_k \in [n]$ are integers such that $s_i \geq \frac{\log^{1.1} n}{p}$ and $\sum_i s_i \leq |U|$.

Then, there exist disjoint subsets $S_1, \dots, S_k \subseteq U$ such that the following holds for every $1 \leq i \leq k$:

(a) $|S_i| = s_i$, and

(b) for every $v \in V(D)$, if $d_D^\pm(v, U) \geq cp|U|$ then

$$d_D^\pm(v, S_i) \geq (1 - \varepsilon)cps_i. \quad (1)$$

Proof. We prove the lemma only for $d^+(v, S_i) \geq (1 - \varepsilon)cps_i$ as the proof for $d^-(v, S_i) \geq (1 - \varepsilon)cps_i$ follows in similar fashion. Let $U = S_1 \cup \dots \cup S_k \cup Z$ be a partition of U taken uniformly at random from all partitions for which $S_i = s_i$ for every $1 \leq i \leq k$ and the “leftover” set Z is of size $|Z| = |U| - \sum_{i=1}^k s_i$. Let $W := \{v \in V(D) \mid d^+(v, U) \geq cp|U|\}$ and let $v \in W$ be an arbitrary vertex from W . The number of out-neighbors of v in S_i is hypergeometrical distributed, thus we have

$$\mathbb{E}[d^+(v, S_i)] = d^+(v, U) \frac{s_i}{|U|} \geq c \log^{1.1} n.$$

Using this and Lemma 2.1 we obtain the following upper bound

$$\Pr \left[d^+(v, S_i) \leq (1 - \varepsilon)d^+(v, U) \frac{s_i}{|U|} \right] \leq e^{-\varepsilon^2 c \log^{1.1} n/2}.$$

The set W has at most n vertices and the size of each part S_i is a positive integer, thus the number of parts k is at most n . Taking the union bound over all parts S_1, \dots, S_k and all vertices in W we get

$$\Pr \left[\exists v \in W \exists i \in [k], d^+(v, S_i) \leq (1 - \varepsilon)d^+(v, U) \frac{s_i}{|U|} \right] \leq n^2 \cdot e^{-\varepsilon^2 c \log^{1.1} n/2} = o(1),$$

which completes the proof. \square

3 Proof of Theorem 1.3

In this section we introduce the definition of an (n, α, p) -pseudorandom digraph, which will be the main object of study throughout the paper. In fact, we prove that for $p = \omega(\frac{\log^8 n}{n})$ and any positive constant α , an (n, α, p) -pseudorandom digraph contains a directed Hamiltonian cycle. Using this result the main theorem follows from a fact that after deleting at most $(1/2 - \beta)$ fraction of the edges from each vertex of $D(n, p)$ the remaining digraph is w.h.p. (n, α, p) -pseudorandom, for some positive constant $\alpha < \beta$.

Definition 3.1. A directed graph D on n vertices is called (n, α, p) -pseudorandom if the following holds:

(P1) for every $v \in V(D)$ we have

$$d_D^\pm(v, V(D)) \geq (1/2 + 2\alpha)np,$$

(P2) for every subset $X \subseteq V(D)$ of size $|X| \leq \frac{\log^2 n}{p}$, we have

$$e_D(X) \leq |X| \log^{2.1} n,$$

(P3) for every two disjoint subsets $X, Y \subseteq V(D)$ of sizes $|X|, |Y| \geq \frac{\log^{1.1} n}{p}$, we have

$$e_D(X, Y) \leq (1 + \alpha/2)|X||Y|p.$$

Intuitively, we require from an (n, α, p) -pseudorandom digraph a certain lower bound on the minimum degree and that it contains no dense subgraph. As it turns out, these properties are sufficient for containing a directed Hamiltonian cycle.

Theorem 3.2. *Let $\alpha > 0$ be a constant and n sufficiently large integer. Then for $p = \omega(\frac{\log^8 n}{n})$, every (n, α, p) -pseudorandom digraph is Hamiltonian.*

Before we give the proof of Theorem 3.2, we first show how it implies Theorem 1.3.

Proof of Theorem 1.3. Let β, n and p be as stated in the theorem and let $\alpha = \beta/4$. By Theorem 3.2, it is sufficient to prove that $G \sim D(n, p)$ w.h.p satisfies that $D = G - H$ is (n, α, p) -pseudorandom, for every $H \subseteq G$ as given in Definition 1.1. Using the fact that in G w.h.p. $d_G^+(v), d_G^-(v) \in (1 \pm o(1))np$ for each $v \in V(G)$, Definition 1.1 implies that we have to show this for all subgraphs $H \subset G$ which satisfy $d_H^+(v), d_H^-(v) \leq (1/2 - \beta)np$, for every $v \in V(H)$. Let us consider one such subgraph H .

First, observe that for every vertex $v \in D$ we have

$$d_D^\pm(v) \geq (1 - o(1) - 1/2 + \beta)np \geq (1/2 + 2\alpha)np, \quad (2)$$

thus the property (P1) holds.

For (P2), let $X \subseteq V(G)$ be an arbitrary subset of size at most $\frac{\log^2 n}{p}$. Since $e_G(X) \sim \text{Bin}(2\binom{|X|}{2}, p)$, by Lemma 2.3 we have

$$\Pr[e_G(X) \geq |X| \log^{2.1} n] \leq \left(\frac{e|X|^2 p}{|X| \log^{2.1} n} \right)^{|X| \log^{2.1} n}.$$

A union bound over the choices of X shows that the probability that there exists a subset $X \subseteq V(G)$ of size $x \leq \frac{\log^2 n}{p}$ such that $e_G(X) \geq |X| \log^{2.1} n$ is at most

$$\begin{aligned} \sum_{x \leq \frac{\log^2 n}{p}} \binom{n}{x} \left(\frac{ex^2 p}{x \log^{2.1} n} \right)^{x \log^{2.1} n} &\leq \sum_{x \leq \frac{\log^2 n}{p}} \left(n \left(\frac{exp}{\log^{2.1} n} \right)^{\log^{2.1} n} \right)^x \\ &\leq \sum_{x \leq \frac{\log^2 n}{p}} \left(n \left(\frac{e}{\log^{0.1} n} \right)^{\log^{2.1} n} \right)^x = o(1). \end{aligned}$$

Hence, (P2) holds in G and therefore in $D \subseteq G$ as well.

The proof of (P3) goes similarly. Consider disjoint subsets $X, Y \subseteq V(G)$ of size at least $\log^{1.1} n/p$. Then by Lemma 2.1 we have

$$\Pr[e_G(X, Y) > (1 + \alpha/2)|X||Y|p] < e^{-\Omega(|X||Y|p)}.$$

A union bound over the choices for X and Y yields that the probability that (P3) fails is upper bounded by

$$\sum_{x, y = \frac{\log^{1.1} n}{p}}^n \binom{n}{x} \binom{n}{y} e^{-\Omega(xyp)} \leq \sum_{x, y = \frac{\log^{1.1} n}{p}}^n n^x n^y e^{-\Omega(\max\{x, y\} \log^{1.1} n)} = o(1).$$

This proves G is w.h.p such that D is an (n, α, p) -pseudorandom digraph, regardless of the choice of H , and thus completes the proof. \square

Throughout the paper, unless stated otherwise, we always assume that D is a (n, α, p) -pseudorandom digraph where α is a positive constant and $p = \omega(\frac{\log^8 n}{n})$. Moreover, with $\{V_1, V_2, V_3, V_4, V_5\}$ we denote the partition of $V(D)$ given by the following claim.

Claim 3.3. *There exists a partition $V(D) = \bigcup_{i=1}^5 V_i$ of the vertices of D , such that the following holds:*

(Q1) *for every $v \in V(D)$ and every $i \in \{1, 2, 3, 4, 5\}$, we have $d_D^\pm(v, V_i) \geq (1/2 + \alpha)|V_i|p$,*

(Q2) *$|V_1| = (1 + o(1))n/\log^3 n$ and $|V_2|, |V_3|, |V_4| \in \left(\frac{\alpha}{5(1+2\alpha)}n, \frac{\alpha}{4(1+2\alpha)}n\right)$.*

Proof. Let $s_1 = (1 + o(1))n/\log^3 n$, $s_2, s_3, s_4 \in (\frac{\alpha}{5(1+2\alpha)}n, \frac{\alpha}{4(1+2\alpha)}n)$ be arbitrarily chosen integers and $s_5 = n - \sum_{i=1}^4 s_i$. As $s_i \geq \frac{\log^{1.1} n}{p}$ for all $i \in \{1, 2, 3, 4, 5\}$, by applying Lemma 2.4 with $c = 1/2 + 2\alpha$, $\varepsilon = \alpha/3$, $V(D)$ (as U), $k = 5$, s_1, s_2, s_3, s_4 and s_5 and by using (P1) we obtain sets V_1, V_2, V_3, V_4 and V_5 (S_1, S_2, S_3, S_4 and S_5 in the Lemma 2.4) such that $V(D) = V_1 \cup V_2 \cup V_3 \cup V_4 \cup V_5$ and

$$d^\pm(v, S_i) \geq (1 - \varepsilon)cs_i p = (1 - \varepsilon) \cdot (1/2 + 2\alpha)s_i p \geq (1/2 + \alpha)s_i p,$$

for every $v \in V(D)$ and $i \in \{1, 2, 3, 4, 5\}$, as required. \square

3.1 Proof of Theorem 3.2

The following two lemmas will serve as our main tool for proving Hamiltonicity of D .

Lemma 3.4 (Absorbing Lemma). *There exists a directed path P^* with $V(P^*) \subseteq V_2 \cup V_3 \cup V_4$ such that for every $W \subseteq V_1$ there is a directed path P_W^* with $V(P_W^*) = V(P^*) \cup W$ and such that P_W^* and P^* have the same endpoints.*

The following lemma states that, under certain assumptions, one can find disjoint σ -walks connecting specified pairs of vertices, for arbitrary σ of length $\Omega(\log n)$. The proof of Connecting Lemma is a modification of a beautiful argument by Montgomery [15].

Lemma 3.5 (Connecting Lemma). *Let ℓ and t be integers such that $\ell \geq 10 \log n$ and $t \geq \frac{4 \log^2 n}{p}$ and let $\{(a_i, b_i)\}_{i=1}^t$ be a family of pairs of vertices from $V(D)$ with $a_i \neq a_j$ and $b_i \neq b_j$ for every distinct $i, j \in [t]$. Assume that $K \subseteq V(D) \setminus \bigcup_{i=1}^t \{a_i, b_i\}$ is such that*

(i) $|K| = \omega(t\ell)$,

(ii) *for every $v \in K$ we have $d^\pm(v, K) \geq (\frac{1}{2} + \alpha)p|K|$ and*

(iii) *for every $i \in [t]$ we have*

$$d^\pm(a_i, K), d^\pm(b_i, K) \geq (1/2 + \alpha)p|K|.$$

Then for every $\sigma \in \{-, +\}^\ell$ there exist t internally disjoint σ -walks P_1, \dots, P_t such that for each i , P_i connects a_i to b_i and $V(P_i) \setminus \{a_i, b_i\} \subseteq K$.

With these two lemmas at hand, we are ready to give a proof of Theorem 3.2.

Proof of Theorem 3.2. Let P^* be a path obtained from Lemma 3.4 and let $U := (V_2 \cup V_3 \cup V_4 \cup V_5) \setminus V(P^*)$. We first show that there exists a family $\{Q_1, \dots, Q_{t'}\}$ of $t' \in [n/\log^5 n, 2n/\log^5 n]$ vertex-disjoint directed paths such that $U = \bigcup_{i=1}^{t'} V(Q_i)$. It follows from property (Q2) that $|V_5| \geq (1 - \frac{\alpha}{1+2\alpha})n$. Furthermore, property (Q1) and $U \subseteq V_2 \cup V_3 \cup V_4 \cup V_5$ imply

$$d^\pm(v, U) \geq d^\pm(v, V_5) \geq (1/2 + \alpha)p|V_5| \geq (1/2 + \alpha)p \frac{1 + \alpha}{1 + 2\alpha} |U| = (1/2 + \alpha/2)p|U|, \quad (3)$$

for every $v \in U$. Applying Lemma 2.4 with $c = 1/2 + \alpha/2$, ε such that $(1 - \varepsilon)c > 1/2 + \alpha/4$, U , $k = \lfloor \log^5 n |U|/n \rfloor$ and $s_i = \lfloor n/\log^5 n \rfloor$ for every $i \in [k]$, together with (3), we obtain disjoint subsets $S_1, \dots, S_k \subseteq U$ such that

$$d^\pm(v, S_i) \geq (1 - \varepsilon)cp|S_i| \geq (1/2 + \alpha')p|S_i|, \quad (4)$$

for some constant $\alpha' > \alpha/4$, every $i \in [k]$ and $v \in U$. We now use the following claim, whose proof we defer to the end of the subsection.

Claim 3.6. *For every $i \in [k - 1]$, there exists a perfect matching from S_i to S_{i+1} .*

Observe that such matchings induce $s := \lfloor n/\log^5 n \rfloor$ vertex-disjoint directed paths $\{Q_1, \dots, Q_s\}$, each of length k , such that $\bigcup_{i=1}^s V(Q_i) = \bigcup_{i=1}^k S_i$. On the other hand, by taking each vertex in $U \setminus \bigcup_{i=1}^k S_i$ to be a 0-length path, we obtain at most s additional paths $\{Q_{s+1}, \dots, Q_{t'}\}$. Note that the family $\{Q_1, \dots, Q_{t'}\}$ satisfies the desired properties.

As a final step, we find a cycle C in D which contains paths $Q_1, \dots, Q_{t'}, P^*$ and maybe some vertices from V_1 . Using Lemma 3.4, we can absorb the remaining vertices from V_1 and obtain a Hamiltonian cycle. We now make this more precise.

For $i \in [t']$, let us denote with a_i and b_i the first and the last vertex on the path Q_i . Furthermore, let $a_{t'+1}$ and $b_{t'+1}$ be the first and the last vertex of the path P^* . Applying Lemma 3.5 with $\ell = 10 \log n$, $t = t' + 1$, the family of pairs $\{(b_i, a_{i+1})\}_{i=1}^{t'} \cup \{(b_{t'+1}, a_1)\}$ and V_1 (as K), we obtain vertex disjoint directed paths $P_1, \dots, P_{t'+1}$. Observe that $Q_1, P_1, Q_2, \dots, Q_{t'}, P_{t'}, P^*, P_{t'+1}$ forms a directed cycle C with $V_2 \cup V_3 \cup V_4 \cup V_5 \subseteq V(C)$. By Lemma 3.4 there is a path $P_{V_1 \setminus V(C)}^*$ with the same endpoints as P^* and such that $V(P_{V_1 \setminus V(C)}^*) = V(P^*) \cup (V_1 \setminus V(C))$. As C contains the path P^* , we can replace P^* with $P_{V_1 \setminus V(C)}^*$, thus obtaining a Hamiltonian cycle. \square

Proof of Claim 3.6. We use the following theorem which is equivalent to Hall's condition (see [19]): There exists a perfect matching from S_i to S_{i+1} if and only if for every subset $X \subseteq S_i$ of size $|X| \leq |S_i|/2$ we have $|N^+(X, S_{i+1})| \geq |X|$ and for every subset $Y \subseteq S_{i+1}$ of size $|Y| \leq |S_{i+1}|/2$ we have $|N^-(Y, S_i)| \geq |Y|$. Assuming the opposite, without loss of generality there exists a subset $X \subseteq S_i$ of size $|X| \leq |S_i|/2$ for which $N^+(X, S_{i+1})$ is contained in a set Y of size exactly $|X| - 1$. We distinguish between two cases:

(a) $|X| \leq \log^2 n / (2p)$. In this case we have that $e_D(X, S_{i+1} \setminus Y) = 0$, and therefore, using (4) we obtain that

$$e_D(X \cup Y) \geq e_D(X, Y) = e_D(X, S_{i+1}) \geq (1/2 + \alpha')|X||S_{i+1}|p = \omega(|X| \log^{2.1} n),$$

which contradicts (P2) (here we use the fact that $|S_{i+1}|p = \omega(\log^{2.1} n)$).

(b) $\log^{1.1} n / p < |X| \leq |S_i|/2$. In this case we have

$$e_D(X, Y) \geq (1/2 + \alpha')|X||S_{i+1}|p = \frac{1 + 2\alpha'}{2}|X||S_{i+1}|p > (1 + \alpha/2)|X||Y|p,$$

which contradicts (P3) (here we use the fact that $\alpha' > \alpha/4$).

The same argument can be applied to a subset $Y \subseteq S_{i+1}$ of size $|Y| \leq |S_{i+1}|/2$. This completes the proof. \square

4 Proof of the Connecting Lemma

In this section we prove Lemma 3.5. The lemma states that if we have a list of t pairs of vertices and a set K of size $\omega(t\ell)$ (where $\ell \geq 10 \log n$) which “behaves” like a random subset of $V(D)$ then for any $\sigma \in \{+, -\}^\ell$ we can connect each pair of vertices via t disjoint σ -walks of length ℓ by using only vertices from K . The proof is obtained by adopting a clever argument due to Richard Montgomery [15] into the setting of resilience. In order to do so, we had to repeat the whole argument.

4.1 Expansion properties and σ -neighborhoods

We start with a lemma which says that for any two (not too small) sets $X, Y \subseteq V(D)$, such that all vertices $x \in X$ have a large degree in Y , X expands to more than a half of vertices in Y .

Lemma 4.1. *Let $X, Y \subseteq V(D)$ be two (not necessarily disjoint) subsets such that $|X| = \lfloor \frac{\log^2 n}{2p} \rfloor$, $|Y| \geq \frac{6 \log^{2.1} n}{\alpha p}$ and for each $x \in X$*

$$d^\pm(x, Y) \geq (1/2 + \alpha/2)p|Y|.$$

Then $|N^+(X, Y)|, |N^-(X, Y)| \geq (1/2 + \alpha/20)|Y|$.

Proof. We only prove that $|N^+(X, Y)| \geq (1/2 + \alpha/20)|Y|$ as the bound on $|N^-(X, Y)|$ can be proven analogously. From property (P2) we have $e(X) \leq |X| \log^{2.1} n$. Now, denote $S_X := N^+(X, Y) \setminus X$. Using the previous inequality together with $d^\pm(x, Y) \geq (1/2 + \alpha/2)p|Y|$ for every $x \in X$ and $|Y|p \geq 6 \log^{2.1} n/\alpha$, we obtain

$$e(X, S_X) \geq e(X, Y) - e(X) \geq (1/2 + \alpha/2)p|X||Y| - |X| \log^{2.1} n \geq (1/2 + \alpha/3)p|X||Y|. \quad (5)$$

Now let us assume $|S_X| < \frac{\log^2 n}{2p}$. We then know that $|X \cup S_X| < \frac{\log^2 n}{p}$. Therefore we can apply property (P2) to the set $X \cup S_X$ and conclude $e(X \cup S_X) \leq |X \cup S_X| \log^{2.1} n \leq 2|X| \log^{2.1} n$. On the other hand, from Equation (5) and the bound on the size of Y we have

$$e(X \cup S_X) \geq e(X, S_X) \geq (1/2 + \alpha/3)p|X||Y| \geq (2 + 3/\alpha)|X| \log^{2.1} n,$$

which is a contradiction.

Next, we assume $\frac{\log^2 n}{2p} \leq |S_X| \leq (1/2 + \alpha/20)|Y|$. Using property (P3) we obtain

$$e(X, S_X) \leq (1 + \alpha/2)(1/2 + \alpha/20)p|X||Y|.$$

Now, combining the previous inequality with (5) we conclude

$$(1/2 + \alpha/3)p|X||Y| \leq e(X, S_X) \leq (1 + \alpha/2)p|X|(1/2 + \alpha/20)|Y|. \quad (6)$$

However, by easy calculation one can check that $(1 + \alpha/2)(1/2 + \alpha/20) < (1/2 + \alpha/3)$, thus Equation (6) gives a contradiction. Therefore we have $|N^+(X, Y)| \geq |S_X| \geq (1/2 + \alpha/20)|Y|$, as required. \square

Next, we introduce the notion of σ -neighborhood. For given sets $A, B \subseteq V(D)$, integer ℓ and $\sigma \in \{+, -\}^\ell$, we define $N^\sigma(A, B)$ as follows,

$$N^\sigma(A, B) := \{x \in B \mid \exists a_x \in A \text{ and a } \sigma\text{-walk } P \text{ connecting } a_x \text{ to } x \text{ and } V(P) \setminus \{a_x\} \subseteq B\}.$$

In the following lemma we show that for two subsets $X, Y \subseteq V(D)$, such that X and Y have good expansion properties, we can find a vertex $x \in X$ which can reach more than a half of the vertices from Y via σ -walks.

Lemma 4.2. *Let $\gamma \in (0, 1)$ be a constant and let ℓ be an integer such that $\ell \geq 2 \log n$. Suppose that $X, Y \subseteq V(D)$ are two disjoint subsets of vertices such that the following holds:*

- (i) $|Y| \geq \frac{6\ell}{\gamma} \cdot \lceil \frac{\log^2 n}{p} \rceil$,
- (ii) $|N^+(X, Y)|, |N^-(X, Y)| \geq \frac{2 \log^2 n}{p}$ and
- (iii) for every subset $S \subseteq Y$ of size $|S| \geq \frac{\log^2 n}{p}$ we have

$$|N^+(S, Y)|, |N^-(S, Y)| \geq (1/2 + \gamma)|Y|.$$

Then for any $\sigma \in \{+, -\}^\ell$ there exists a vertex $x \in X$ such that

$$|N^\sigma(x, Y)| \geq (1/2 + \gamma/2)|Y|.$$

Proof. Recall that $\sigma^i = (\sigma(1), \dots, \sigma(i))$. We first show that there exists a vertex $x \in X$ such that $|N^{\sigma^{\ell-1}}(x, Y)| \geq \frac{2 \log^2 n}{p}$. In order to do so, we make use of the following claim.

Claim 4.3. *Let $i < \ell$ be an integer and $A \subseteq X$ such that $|N^{\sigma^i}(A, Y)| \geq \frac{2 \log^2 n}{p}$. Then there exists a subset $A' \subseteq A$ such that $|A'| \leq \lceil |A|/2 \rceil$ and*

$$|N^{\sigma^{i+1}}(A', Y)| \geq \frac{2 \log^2 n}{p}.$$

By assumption (ii) we have $|N^{\sigma(1)}(X, Y)| \geq 2 \log^2 n / p$, thus applying Claim 4.3 repeatedly $\ell - 2$ times we obtain a set $X' \subseteq X$ such that $|X'| \leq \lceil |X|/2^{\ell-2} \rceil$ and $|N^{\sigma^i}(X', Y)| \geq \frac{2 \log^2 n}{p}$. Since $|X| \leq n$ and $\ell - 2 \geq \log n$, it follows that $|X|/2^{\ell-2} \leq 1$ and therefore $|X'| = 1$. Hence, there exists $x \in X$ such that $|N^{\sigma^{\ell-1}}(x, Y)| \geq \frac{2 \log^2 n}{p}$.

Let now $M \subseteq N^{\sigma^{\ell-1}}(x, Y)$ be a subset of size $|M| = \lceil \frac{\log^2 n}{p} \rceil$ and note that, by definition, for each $w \in M$ there exists a $\sigma^{\ell-1}$ -walk P_w connecting x to w with $V(P_w) \setminus \{x\} \subseteq Y$. Let $V^* := (\bigcup_{w \in M} V(P_w)) \setminus \{x\}$. Using assumptions (ii) and (iii) we have

$$|N^{\sigma(\ell)}(M, Y \setminus V^*)| \geq (1/2 + \gamma)|Y| - \ell|M| \geq (1/2 + \gamma/2)|Y|.$$

Observe that $N^{\sigma(\ell)}(M, Y \setminus V^*) \subseteq N^\sigma(x, Y)$ and therefore $|N^\sigma(x, Y)| \geq (1/2 + \gamma/2)|Y|$.

In order to complete the proof it remains to prove Claim 4.3.

Proof of Claim 4.3. First, note that there exists a subset $A' \subseteq A$ such that $|A'| \leq \lceil |A|/2 \rceil$ and $|N^{\sigma^i}(A', Y)| \geq \frac{\log^2 n}{p}$. Indeed, this is true as otherwise taking an arbitrary partition of the set $A = S \cup T$, such that $|S|, |T| \leq \lceil |A|/2 \rceil$, yields

$$|N^{\sigma^i}(A, Y)| \leq |N^{\sigma^i}(S, Y)| + |N^{\sigma^i}(T, Y)| < \frac{2\log^2 n}{p},$$

which contradicts the assumption that $|N^{\sigma^i}(A, Y)| \geq \frac{2\log^2 n}{p}$.

Let $H \subseteq N^{\sigma^i}(A', Y)$ be an arbitrary subset of size $|H| = \lceil \frac{\log^2 n}{p} \rceil$. Using assumption (iii) we have $|N^{\sigma^{(i+1)}}(H, Y)| \geq (1/2 + \gamma)|Y|$. We know that for each $v \in H$ there exist a σ^i -walk P_v connecting a vertex from A' to the vertex v . Let us denote $V^* := \cup_{v \in H} V(P_v)$. Using the upper bound on i we have $|V^*| \leq \ell|H|$ and thus

$$|N^{\sigma^{(i+1)}}(H, Y \setminus V^*)| \geq (1/2 + \gamma)|Y| - \ell|H| \geq \frac{2\log^2 n}{p},$$

where the second inequality follows from assumption (i). Finally, observe that $N^{\sigma^{(i+1)}}(H, Y \setminus V^*) \subseteq N^{\sigma^{i+1}}(A', Y)$ and hence we have $|N^{\sigma^{i+1}}(A', Y)| \geq \frac{2\log^2 n}{p}$. \square

This completes the proof of the lemma. \square

4.2 The proof

The following lemma is an approximate version of the Connecting Lemma and it is used as the main building block in the proof of the Connecting Lemma. Namely, the lemma states that for a given set of pairs $\{(a_i, b_i)\}_{i=1}^t$, sets R_A, R_B with good expansion properties we can connect half of the pairs via long σ -walks using only vertices from $R_A \cup R_B$.

Lemma 4.4. *Let $\gamma \in (0, 1)$ be a constant, let ℓ and t be integers such that $\ell \geq 5 \log n$ and $t \geq \frac{4\log^2 n}{p}$ and let $\{(a_i, b_i)\}_{i=1}^t$ be a family of pairs of vertices from $V(D)$ with $a_i \neq a_j$ and $b_i \neq b_j$ for every distinct $i, j \in [t]$. Furthermore, let $R_A, R_B \subseteq V(D) \setminus (\cup_{i=1}^t \{a_i, b_i\})$ be disjoint subsets such that the following holds:*

(i) $|R_A|, |R_B| \geq 12t\ell/\gamma$ and

(ii) for $X \in \{A, B\}$ and for every set $S \subseteq R_A \cup R_B \cup \cup_{i=1}^t \{a_i, b_i\}$ of size at least $\frac{\log^2 n}{p}$ we have

$$|N^+(S, R_X)|, |N^-(S, R_X)| \geq (1/2 + \gamma)|R_X|.$$

Then for any $\sigma \in \{+, -\}^\ell$ there exists a subset of indices $\mathcal{I} \subseteq [t]$ of size $s := \lfloor t/2 \rfloor$ and s internally disjoint σ -walks P_i which connect a_i to b_i (where $i \in \mathcal{I}$), and such that $V(P_i) \setminus \{a_i, b_i\} \subseteq R_A \cup R_B$.

Proof. We prove the existence of set \mathcal{I} of size $\lfloor t/2 \rfloor$ and the required σ -walks by induction. Assume that there exists $\mathcal{I} = \{i_1, \dots, i_{s'}\} \subseteq [t]$ with $s' < \lfloor t/2 \rfloor$, and s' vertex-disjoint σ -walks P_i , connecting a_i to b_i , where $i \in \mathcal{I}$. Let us define

$$R'_A := R_A \setminus \cup_{i \in \mathcal{I}} V(P_i), \quad R'_B := R_B \setminus \cup_{i \in \mathcal{I}} V(P_i) \quad \text{and} \quad \mathcal{I}' := [t] \setminus \mathcal{I}.$$

Next, we show how to find a σ -walk P connecting some a_i to b_i where $i \in \mathcal{I}'$ such that $V(P) \setminus \{a_i, b_i\} \subseteq R'_A \cup R'_B$. Let h_A and h_B be two integers such that $h_A, h_B \geq 2 \log n$ and $h_A + h_B + 1 = \ell$, and consider σ^{h_A} and $\bar{\sigma}^{h_B}$ (recall that $\sigma^{h_A} = (\sigma(1), \dots, \sigma(h_A))$ and $\bar{\sigma}^{h_B} = (\bar{\sigma}(\ell), \dots, \bar{\sigma}(\ell - h_B + 1))$). We make use of the following claim.

Claim 4.5. *There exists an index $i \in \mathcal{I}'$ for which the following holds:*

$$|N^{\sigma^{h_A}}(a_i, R'_A)| \geq (1/2 + \gamma/4)|R'_A|$$

and

$$|N^{\bar{\sigma}^{h_B}}(b_i, R'_B)| \geq (1/2 + \gamma/4)|R'_B|.$$

Before we prove this claim we show how to finish the proof of the lemma. Let (a_i, b_i) be a pair of vertices with index obtained by Claim 4.5. For $S := N^{\sigma^{h_A}}(a_i, R'_A)$ we have that $|S| \geq (1/2 + \gamma/4)|R'_A|$. As R' is obtained by removing vertices of at most $|\mathcal{I}|$ many σ -walks we know $|R'_A| \geq |R_A| - t\ell \geq 11t\ell$, and thus $|S| \geq \log^2 n/p$. By assumption (ii) we have that $N^{\sigma(h_A+1)}(S, R_B) \geq (1/2 + \gamma)|R_B|$ and consequently

$$N^{\sigma(h_A+1)}(S, R'_B) \geq (1/2 + \gamma)|R_B| - t\ell \stackrel{(i)}{\geq} (1/2 + \gamma/2)|R_B| \geq (1/2 + \gamma/2)|R'_B|. \quad (7)$$

On the other hand we know from Claim 4.5 that $|N^{\bar{\sigma}^{h_B}}(b_i, R'_B)| \geq (1/2 + \gamma/4)|R'_B|$. This implies together with Equation (7) that there exist $v \in N^{\sigma^{h_A}}(a_i, R'_A)$ and $w \in N^{\bar{\sigma}^{h_B}}(b_i, R'_B)$ such that

$$w \in N^{\sigma(h_A+1)}(v).$$

Therefore we can construct a σ -walk P connecting a_i to b_i such that P is vertex disjoint from all previous σ -walks. Now it only remains to prove Claim 4.5.

Proof of Claim 4.5. The idea of the proof is to repeatedly apply Lemma 4.2. First, we apply Lemma 4.2 to $X := \bigcup_{i \in \mathcal{I}'} \{a_i\}$, $Y := R'_A$, $\gamma/2$ (as γ) and obtain a vertex $v_1 \in X$ such that $|N^{\sigma^{h_A}}(a, R'_A)| \geq (1/2 + \gamma/4)|R'_A|$. Next, we apply the lemma again but now to $X := X \setminus \{v_1\}$ instead (with other parameters unchanged) and obtain $v_2 \in X \setminus \{v_1\}$. After k steps of this procedure we obtain vertices $\{v_1, \dots, v_k\}$ with the property $|N^{\sigma^{h_A}}(v_i, R'_A)| \geq (1/2 + \gamma/4)|R'_A|$, for all $1 \leq i \leq k$.

Let us now argue that we can indeed apply Lemma 4.2 and, moreover, estimate the number of steps k . Note that the condition (i) from Lemma 4.2 is satisfied as

$$|R'_A| \geq 12t\ell/\gamma \geq \frac{20 \log^3 n}{\gamma p}.$$

On the other hand, using property (ii) of Lemma 4.4 and the fact that $|R_A| - |R'_A| \leq t\ell$ we get

$$|N^-(S, R'_A)|, |N^+(S, R'_A)| \geq (1/2 + \gamma)|R_A| - t\ell \geq (1/2 + \gamma/2)|R'_A|,$$

for any $S \subseteq R'_A \cup \bigcup_{i \in \mathcal{I}'} \{a_i\}$ of size at least $\frac{\log^2 n}{p}$. Therefore, $X = \bigcup_{i \in \mathcal{I}'} \{a_i\} \setminus \{v_1, \dots, v_i\}$ satisfies assumption (ii) of Lemma 4.2 as long as $|\mathcal{I}'| - i > \log^2 n/p$. This implies that we can iterate the process for at least $k \geq |\mathcal{I}'|/2 + 1$ steps as $|\mathcal{I}'| > t/2 \geq \frac{2 \log^2 n}{p}$. Thus, we obtain $V_A := \{v_1, \dots, v_k\}$ with $v_j \in \{a_i\}_{i \in \mathcal{I}'}$ and

$$|N^{\sigma^{h_A}}(v_j, R'_A)| \geq (1/2 + \gamma/4)|R'_A|$$

for all $j \in [k]$.

By using the analogous argument with $\{b_i\}_{i \in \mathcal{I}'}$ and R'_B we obtain $V_B := \{w_1, \dots, w_k\}$ such that $k > |\mathcal{I}'|/2$ and $w_j \in \{b_i\}_{i \in \mathcal{I}'}$ with the property

$$|N^{\bar{\sigma}^{h_B}}(w_j, R'_B)| \geq (1/2 + \gamma/4)|R'_B|$$

for all $j \in [k]$. Therefore, there must exist $i \in \mathcal{I}'$ such that $a_i \in V_A$ and $b_i \in V_B$, as required by the claim. \square

This finishes the proof of Lemma 4.4. \square

Before proving Connecting Lemma we need to introduce the following definitions.

Definition 4.6. Let T be a rooted tree with edges oriented arbitrarily. Let $\mathcal{L}(T)$ denote the set of leaves of T and let $\sigma \in \{+, -\}^\ell$ for some integer ℓ . We say that T is a σ -tree if for each $v \in \mathcal{L}(T)$ the unique path from the root of T to v is a σ -walk.

Definition 4.7. Let $\tau \in \{+, -\}$ and let $X, Y \subseteq V(D)$ be two disjoint sets. We say that there is a $(2, \tau)$ -matching between X and Y that saturates X if for each $x \in X$ there are two distinct vertices $y_x^1, y_x^2 \in Y$ such that $\{y_x^1, y_x^2\} \in N^\tau(x)$ and $\{y_x^1, y_x^2\} \cap \{y_{x'}^1, y_{x'}^2\} = \emptyset$ for $x \neq x'$.

We are finally ready to prove the main lemma of this section.

Proof of Lemma 3.5. Let σ be an arbitrary element of $\{+, -\}^\ell$ and let $\varepsilon > 0$ be a sufficiently small constant (to be determined later). Throughout the proof we make use of the following parameters:

$$\begin{aligned} h &= \max \left\{ \frac{6 \log^{2.2} n}{p}, 2t \right\}, \\ m &= \lceil \log_2 t \rceil + 1, \\ s_i &= h \quad \text{for } 1 \leq i \leq 2m, \\ s_{2m+1} &= s_{2m+2} = \frac{|K|}{4}, \\ k &= 2m + 2. \end{aligned}$$

Applying Lemma 2.4 to s_1, s_2, \dots, s_k , ε , $1/2 + \alpha$ (as γ), p , K (as U) and D we obtain disjoint subsets $S_1, \dots, S_k \subseteq K$, such that for every $1 \leq i \leq k$ the following holds:

(a) $|S_i| = s_i$, and

(b) for every $v \in V(D)$, if $d_D^\pm(v, K) \geq (1/2 + \alpha)p|K|$, then

$$d_D^\pm(v, S_i) \geq (1 - \varepsilon)(1/2 + \alpha)ps_i. \quad (8)$$

Using (8), properties (ii) and (iii) and the fact that ε is sufficiently small, we obtain that for any $v \in \{a_i \cup b_i\}_{i=1}^t \cup K$ and for any set S_i the following holds:

$$d_D^\pm(v, S_i) \geq (1/2 + \alpha/2)ps_i. \quad (9)$$

For simplicity of presentation, let us denote $A_0 := \bigcup_{i=1}^t \{a_i\}$, $B_0 := \bigcup_{i=1}^t \{b_i\}$, $A_i := S_i$ for each $1 \leq i \leq m$, $B_i := S_{m+i}$ for each $1 \leq i \leq m$, $R_A := S_{2m+1}$ and $R_B := S_{2m+2}$.

We first describe, informally, the strategy for finding σ -walks. In a first step, we apply Lemma 4.4 to find $t/2$ σ -walks between vertices in A_0 and B_0 . Then we find a $(2, \sigma(1))$ -matching between the leftovers in A_0 and the vertices in A_1 and a $(2, \bar{\sigma}(\ell))$ -matching between the leftovers in B_0 and the vertices in B_1 . Let A'_1 denote the set of vertices that are matched to a leftover of A_0 and analogously define B'_1 . Observe that $|A'_1| = |B'_1| \geq t$ and therefore one can apply Lemma 4.4 to find $|A'_1|/2$ vertex disjoint κ -walks between vertices in A'_1 and B'_1 , where $\kappa := (\sigma(2), \dots, \sigma(\ell - 1))$. Note that extending the walks with the matchings yields σ -walks between at least $t/4$ leftovers of A_0 and the corresponding leftovers of B_0 . By iteratively continuing this process for roughly $\log t$ steps, we construct all the desired walks.

Before proceeding with the description of the procedure, we define the following invariant which we maintain in every step $0 \leq s \leq m$:

(X1) $\mathcal{I}_s \subseteq [t]$ for which

- (a) if $s < m$ then $|\mathcal{I}_s| = t - \lceil t/2^s \rceil$
- (b) if $s = m$ then $|\mathcal{I}_s| = t$

(X2) $\mathcal{P}_s = \{P_i\}_{i \in \mathcal{I}_s}$ is a collection of vertex-disjoint σ -walks such that P_i connects a_i to b_i and $V(P_i) \setminus \{a_i, b_i\} \subseteq (\bigcup_{k=1}^s (A_k \cup B_k)) \cup R_A \cup R_B$, for all $i \in \mathcal{I}_s$

(X3) $\mathcal{T}_A^s = \{T_A^i\}_{i \in [t] \setminus \mathcal{I}_s}$ is a collection of σ^s -trees and $\mathcal{T}_B^s = \{T_B^i\}_{i \in [t] \setminus \mathcal{I}_s}$ is a collection of $\bar{\sigma}^s$ -trees such that for each $i \in [t] \setminus \mathcal{I}_s$

- (a) T_A^i is rooted at a_i , $\mathcal{L}(T_A^i) \subseteq A_s$, $|\mathcal{L}(T_A^i)| = 2^s$ and $V(T_A^i) \setminus \{a_i\} \subseteq \bigcup_{j=1}^s A_j$
- (b) T_B^i is rooted at b_i , $\mathcal{L}(T_B^i) \subseteq B_s$, $|\mathcal{L}(T_B^i)| = 2^s$ and $V(T_B^i) \setminus \{b_i\} \subseteq \bigcup_{j=1}^s B_j$

(X4) for any two $T_1, T_2 \in \mathcal{T}_A^s \cup \mathcal{T}_B^s$ rooted at v_1 and v_2 we have $(V(T_1) \setminus v_1) \cap (V(T_2) \setminus v_2) = \emptyset$

(X5) for any two $T \in \mathcal{T}_A^s \cup \mathcal{T}_B^s$ and any $P \in \mathcal{P}_s$ we have $V(T) \cap V(P) = \emptyset$

The set \mathcal{I}_s represents a set of indices of pairs which are connected by a σ -walk up to step s . The collection \mathcal{P}_s contains σ -walks created up to step s between pairs with indices in \mathcal{I}_s and the \mathcal{T}_A^s and \mathcal{T}_B^s are collections of trees for each element of a pair not connected by a σ -walk up to step s .

First, the invariant clearly holds for $s = 0$, $\mathcal{I}_0 = \emptyset$, $\mathcal{T}_A^0 = A_0$, $\mathcal{T}_B^0 = B_0$ and $\mathcal{P}_0 = \emptyset$. Suppose that the invariant holds for some s such that $s < m$, we will show how to extend it to $s + 1$. Denote $A'_s = \bigcup_{T \in \mathcal{T}_A^s} \mathcal{L}(T)$ and $B'_s = \bigcup_{T \in \mathcal{T}_B^s} \mathcal{L}(T)$. Let $\{a'_i, b'_i\}_{i=1}^r$ be a perfect matching between vertices of A'_s and B'_s with the following property: for every $1 \leq i \leq r$ there is $j \in [t]$ such that a'_i and b'_i are leaves of trees rooted at a_j and b_j . Using (X1) and (X3) we obtain that $r = 2^s(t - |\mathcal{I}_s|) \geq 2^s \lceil t/2^s \rceil \geq t$. Next, let $R'_A = R_A \setminus \bigcup_{i \in \mathcal{I}_s} V(P_i)$ and let $R'_B = R_B \setminus \bigcup_{i \in \mathcal{I}_s} V(P_i)$. Using Claim 4.9 below it follows that for every $X \in \{A, B\}$ and every subset $S \subseteq K$ such that $|S| \geq \frac{\log^2 n}{p}$ we have

$$|N^+(S, R'_X)|, |N^-(S, R'_X)| \geq (1/2 + \alpha/40)|R'_X|.$$

Therefore, we can apply Lemma 4.4 to $\alpha/40$ (as γ), the family of pairs $\{(a'_i, b'_i)\}_{i=1}^r$, R'_A (as R_A), R'_B (as R_B) and $\kappa := (\sigma(s+1), \dots, \sigma(\ell-s))$, and obtain the following: a set of indices $\mathcal{J} \subseteq [r]$ of size $|\mathcal{J}| = \lfloor r/2 \rfloor = 2^{s-1}(t - |\mathcal{I}_s|)$ and a collection of vertex-disjoint κ -walks W_i , such that a path W_i connects a'_i to b'_i and $V(W_i) \setminus \{a'_i, b'_i\} \subseteq R'_A \cup R'_B$, for each $i \in \mathcal{J}$.

Let us pick a subset $\mathcal{J}' \subseteq \mathcal{J}$ of size $\lfloor |\mathcal{J}|/2^s \rfloor = \lfloor (t - |\mathcal{I}_s|)/2 \rfloor$ such that each tree from $\mathcal{T}_B^s \cup \mathcal{T}_A^s$ has at most one leaf indexed with some $i' \in \mathcal{J}'$. Note that this is possible since each tree has 2^s leaves. For technical reasons, when $s = m - 1$ we pick \mathcal{J}' of size $|\mathcal{J}'| = 1$. Now, for any $j \in \mathcal{J}'$ let Q_j^1 be the unique σ^s -walk in the tree $T_1 \in \mathcal{T}_A^s$ containing a'_j which connects the root of T_1 to a'_j . Similarly, let Q_j^2 be the unique $\bar{\sigma}^s$ -walk in the tree $T_2 \in \mathcal{T}_B^s$ containing b'_j which connects the root of T_2 to b'_j . We know that Q_j^1 and Q_j^2 start in vertices with the same index from A_0 and B_0 , by the definition of matching between A'_s and B'_s . Combining the paths Q_j^1, Q_j^2 and W_j we obtain a σ -walk P_{i_j} which connects vertex a_{i_j} to b_{i_j} for some $i_j \in [t] \setminus \mathcal{I}_s$. We define $\mathcal{P}_{s+1} = \mathcal{P}_s \cup (\bigcup_{j \in \mathcal{J}'} P_{i_j})$ and $\mathcal{I}_{s+1} = \mathcal{I}_s \cup \{i_j \mid j \in \mathcal{J}'\}$. Using the fact that $|\mathcal{J}'| = \lfloor (t - |\mathcal{I}_s|)/2 \rfloor$ when $s < m - 1$ we obtain

$$|\mathcal{I}_{s+1}| = |\mathcal{I}_s| + \lfloor (t - |\mathcal{I}_s|)/2 \rfloor = t - (t - |\mathcal{I}_s| - \lfloor (t - |\mathcal{I}_s|)/2 \rfloor) = t - \lceil (t - |\mathcal{I}_s|)/2 \rceil \stackrel{(X1)}{=} t - \lceil t/2^{s+1} \rceil. \quad (10)$$

If $s = m - 1$ it follows from the invariant that $|\mathcal{I}_s| = t - 1$ and therefore that $|\mathcal{I}_{s+1}| = t$. Note that the invariant (X2) holds directly by the construction of the new σ -walk P_{i_j} .

In order to show (X3) let $\mathcal{T}'_A \subseteq \mathcal{T}_A^s$ and $\mathcal{T}'_B \subseteq \mathcal{T}_B^s$ be the subsets which contain all trees that are rooted at vertices from $\cup_{i \in \mathcal{I}_{s+1}} \{a_i, b_i\}$. For $L_A = \cup_{T \in \mathcal{T}'_A} \mathcal{L}(T)$ and $L_B = \cup_{T \in \mathcal{T}'_B} \mathcal{L}(T)$ it follows from (10) that $|L_A| = |L_B| = 2^s \cdot \lceil t/2^{s+1} \rceil$. Using Claim 4.8 below we conclude that there exist an L_A -saturating $(2, \sigma(s+1))$ -matching M_A from L_A to A_{s+1} , and a L_B -saturating $(2, \bar{\sigma}(\ell-s))$ -matching M_B from L_B to B_{s+1} . For every σ^j -tree $T \in \mathcal{T}_A^s$ we denote by T^+ the σ^j -tree obtained by extending T with the arcs of the matching M_A incident to $\mathcal{L}(T)$. Similarly, for every $\bar{\sigma}^j$ -tree $T \in \mathcal{T}_B^s$ we denote by T^+ the $\bar{\sigma}^j$ -tree obtained by extending T with the arcs of the matching M_B incident to $\mathcal{L}(T)$. Finally, let $\mathcal{T}_A^{s+1} = \cup_{T \in \mathcal{T}'_A} T^+$ and let $\mathcal{T}_B^{s+1} = \cup_{T \in \mathcal{T}'_B} T^+$. It follows from our construction that \mathcal{T}_A^{s+1} and \mathcal{T}_B^{s+1} satisfy (X3) and (X4). This completes the proof of the lemma.

Claim 4.8. *For every $0 \leq i \leq m - 1$, every $\tau \in \{+, -\}$ and $X \in \{A, B\}$ the following holds. For every subset $S \subseteq X_i$ of size $|S| \leq |X_{i+1}|/8$ there is a $(2, \tau)$ -matching from S to X_{i+1} that saturates S .*

Proof of Claim 4.8. We prove the existence of a $(2, +)$ -matching from S to X_{i+1} that saturates S and the proof for such a $(2, -)$ -matching follows similarly. Using Hall's Theorem (see e.g. [19]), it is sufficient to prove that for every $S' \subseteq S$ it holds $|N^+(S', X_{i+1})| \geq 2|S'|$. Assume the existence of a subset $S' \subseteq S$ that violates Hall's condition, i.e. $|N^+(S', X_{i+1})| < 2|S'|$. If $|S'| \leq \frac{\log^2 n}{3p}$ then, since $|S' \cup N^+(S', X_{i+1})| \leq \frac{\log^2 n}{p}$, by property (P2) we obtain that

$$e_D(S' \cup N^+(S', X_{i+1})) \leq 3|S'| \log^{2.1} n. \quad (11)$$

Moreover, it follows from (9) that $d^\pm(s', X_{i+1}) \geq (1/2 + \alpha/2)p|X_{i+1}|$, for all $s' \in S'$. All in all, we get

$$e_D(S', N^+(S', X_{i+1})) \geq (1/2 + \alpha/2)p|S'| |X_{i+1}| > 3|S'| \log^{2.2} n, \quad (12)$$

where the second inequality follow from $|X_{i+1}| \geq \frac{6 \log^{2.2} n}{p}$. The last inequality together with (11) leads to a contradiction.

If on the other hand $|S'| > \frac{\log^2 n}{3p}$ and $|N^+(S', X_{i+1})| < 2|S'|$, it follows from (P3) and the assumption on S' that

$$e_D(S', N^+(S', X_{i+1})) \leq (1 + \alpha/2)2p|S'|^2.$$

However by (12) and the assumption $|X_{i+1}| \geq 4|S'|$ we have $e_D(S', N^+(S', X_{i+1})) \geq (2 + 2\alpha)p|S'|^2$. Therefore, by combining the previous inequalities we obtain

$$(2 + \alpha)p|S'|^2 \geq e_D(S', N^+(S', X_{i+1})) \geq (2 + 2\alpha)p|S'|^2,$$

which is a contradiction. \square

Claim 4.9. *For every $1 \leq s \leq m - 1$ and every subset $S \subseteq K$ such that $|S| \geq \frac{\log^2 n}{p}$ and $X \in \{A, B\}$ the following holds:*

$$|N^+(S, R'_X)|, |N^-(S, R'_X)| \geq (1/2 + \alpha/40)|R'_X|, \quad (13)$$

where $R'_X := R_X \setminus \cup_{P \in \mathcal{P}_s} V(P)$.

Proof of Claim 4.9. We prove the claim for $|N^+(S, R'_X)|$ as the proof for $|N^-(S, R'_X)|$ follows analogously. By Equation (9) we know that for $X \in \{A, B\}$, for any $v \in \{a_i, b_i\}_{i=1}^t \cup K$ the following holds at each step s :

$$d^\pm(v, R_X) \geq (1/2 + \alpha/2)p|R_X|.$$

Applying Lemma 4.1 to S (as X), R_X (as Y) we get $|N^+(S, R_X)| \geq (1/2 + \alpha/20)|R_X|$. Since $|\cup_{P \in \mathcal{P}_s} V(P)| \leq t\ell$ and $|R_X| = \omega(t\ell)$ we conclude that

$$|N^+(S, R'_X)| \geq (1/2 + \alpha/20)|R_X| - t\ell \geq (1/2 + \alpha/40)|R'_X|.$$

□

□

5 Proof of the Absorbing Lemma

In this section we prove Lemma 3.4. A main ingredient in our proof is the concept of an *absorber*. Roughly speaking, in our setting (that is, finding a directed Hamilton cycle in a digraph D) an absorber A_x for a vertex x is a digraph which contains x and two designated vertices s_x and t_x , such that A_x contains two $s_x t_x$ -paths: one which consists of all vertices in $V(A_x)$ and the other which consists of all vertices in $V(A_x) \setminus \{x\}$.

Definition 5.1. Let ℓ_x be an integer and A_x a digraph of size $\ell_x + 1$. Then for some distinct vertices $x, s_x, t_x \in V(A_x)$, the digraph A_x is called an absorber for a vertex x with starting point s_x and a terminal point t_x , if it contains an $s_x t_x$ -path P_x , referred to as the non-absorbing path, of length $\ell_x - 1$ that does not contain x , and an $s_x t_x$ -path P'_x of length ℓ_x which is referred to as the absorbing path.

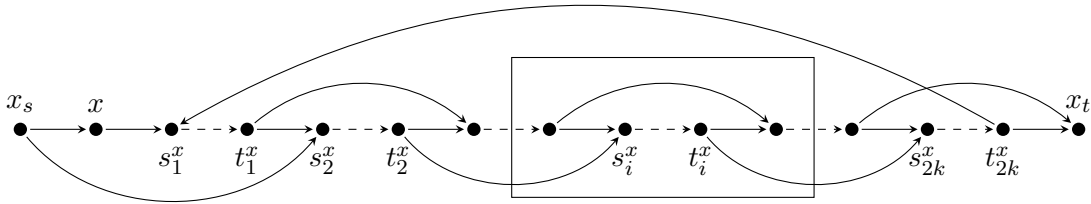


Figure 1: The absorber for $k = 3$. The cycle C is drawn with solid arrows. The dashed arrows represent directed paths of arbitrary length. The part inside the rectangle can be repeated to obtain absorbers for larger k .

In the following lemma we describe the structure of our absorber.

Lemma 5.2. Let k and ℓ be integers and consider a digraph A_x of size $3 + 2k(\ell + 1)$ constructed as follows:

- (i) A_x consists of a cycle C of length $4k + 3$ with an orientation of the edges and labeling of the vertices as shown in Figure 1, and
- (ii) A_x contains $2k$ pairwise disjoint directed $s_i^x t_i^x$ -paths P_i (for each $i \in [2k]$), each of which is of length ℓ .

Then A_x is an absorber for the vertex x .

Proof. It is easy to see that

$$P'_x := x_s, x, s_1^x, P_1, t_1^x, s_2^x, P_2 \dots, t_{2k}^x, x_t$$

is an absorbing path. On the other hand, the path

$$P_x := x_s, s_2^x, P_2, t_2^x, s_4^x, P_4, \dots, s_{2k}^x, t_{2k}^x, s_1^x, P_1, t_1^x, s_3^x, P_3, \dots, t_{2k-1}^x, x_t$$

uses all vertices except x , thus it is a non-absorbing path. We refer the reader to Figure 1 for clarification. \square

The proof of the Absorbing Lemma consists of two main steps. First, we show how to construct an absorber A_x for each $x \in V_1$ such that the non-absorbing path of A_x is contained in $V_2 \cup V_3$ and $V(A_x) \cap V(A'_x) = \emptyset$ for $x \neq x'$. Second, using Lemma 3.5 we connect non-absorbing paths of each absorber into one long path using vertices from V_4 .

We build the absorbers A_x in D by first finding the cycle of the absorber and then connecting all the designated pairs of vertices via directed paths. To do so we use Lemma 3.5 (note that a cycle is a σ -walk, for some σ).

Proof of Lemma 3.4. Let $k := 3\lceil \log n \rceil$ and let A_x be an absorber given by Lemma 5.2. Recall that A_x contains a cycle C_x of length $4k + 3 = 12\lceil \log n \rceil + 3$ with a prescribed orientation σ and $2k = 6\lceil \log n \rceil$ disjoint directed paths P_1, \dots, P_n , of length $\ell := 10\lceil \log n \rceil$, connecting the pairs of designated vertices $(s_1^x, t_1^x), \dots, (s_{2k}^x, t_{2k}^x)$ on the cycle. In order to find for each $x \in V_1$ such a cycle, we apply Lemma 3.5 to the set V_2 (as K), $\ell = 12\lceil \log n \rceil + 3$, $t = |V_1|$ and a family of pairs $\{(x, x)\}_{x \in V_1}$ and thereby obtain for every $x \in V_1$ a σ -walk of length $4k + 3$ from x to itself, which is a cycle C_x as required for the absorber. Moreover, all obtained cycles $\{C_x\}_{x \in V_1}$ are disjoint and contain (apart from the absorbing vertices) only vertices in V_2 . Note that we can apply Lemma 3.5 as $V_2 = \omega(|V_1| \log n)$, $t = |V_1| \geq \frac{4\log^2 n}{p}$ and by property (Q1) we have that (ii) and (iii) from Lemma 3.5 are true.

Next, using the vertices in V_3 , for each $x \in V_1$ we connect the pair of designated vertices (s_i^x, t_i^x) on a cycle C_x by a directed path. For this aim we apply Lemma 3.5 to V_3 (as K) with $\ell := 10\lceil \log n \rceil$, $t = 2k|V_1|$, and $\{(s_i^x, t_i^x) \mid x \in V_1, 1 \leq i \leq 2k\}$ to find all the required paths to complete the absorbers. We can indeed do that as $|V_3| = \omega(|V_1| \log^2 n)$, $t = 2k|V_1| \geq \frac{4\log^2 n}{p}$ and by property (Q1) we have that (ii) and (iii) from Lemma 3.5 are true.

Finally, we build a directed path which contains all the non-absorbing paths of the absorbers. To do so, recall that by Definition 5.1 every absorber A_x has a start vertex s_x and a terminal vertex t_x . Let us arbitrarily enumerate vertices from V_1 as $V_1 = \{x_1, \dots, x_h\}$, where $h = |V_1|$. Apply Lemma 3.5 to V_4 (as K), $\ell = 10\lceil \log n \rceil$, $t = |V_1| - 1$ and a family of pairs $\{(t_{x_i}, s_{x_{i+1}})\}_{i \in [h-1]}$ to find the required paths of length ℓ that connect all non-absorbing paths of the absorbers into one directed path P^* . Again, we are allowed to apply the lemma as $|V_4| = \omega(|V_1| \log n)$ and by property (Q1) we have that (ii) and (iii) from Lemma 3.5 are true.

It is now easy to see that the path P^* has the required properties. Let $W \subseteq V_1$ be an arbitrary subset of V_1 and let $\{A_w \mid w \in W\}$ be the set of absorbers for vertices in W . By the definition of absorber for each A_w there is an absorbing path starting and ending at the same vertices as the non-absorbing path, but which contains vertex w as well. By replacing non-absorbing paths of $\{A_w \mid w \in W\}$ in P^* with corresponding absorbing paths we obtain a path P_W^* which has the same endpoints as P^* and $V(P_W^*) = V(P^*) \cup W$. \square

References

- [1] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2004.
- [2] J. Balogh, B. Csaba, and W. Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011.
- [3] S. Ben-Shimon, M. Krivelevich, and B. Sudakov. Local resilience and hamiltonicity maker–breaker games in random regular graphs. *Combinatorics, Probability and Computing*, 20(02):173–211, 2011.
- [4] S. Ben-Shimon, M. Krivelevich, and B. Sudakov. On the resilience of hamiltonicity and optimal packing of hamilton cycles in random graphs. *SIAM Journal on Discrete Mathematics*, 25(3):1176–1193, 2011.
- [5] B. Bollobás. *Random graphs*. Springer, 1998.
- [6] J. Böttcher, Y. Kohayakawa, and A. Taraz. Almost spanning subgraphs of random graphs after adversarial edge removal. *Electronic Notes in Discrete Mathematics*, 35:335–340, 2009.
- [7] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.
- [8] A. Frieze and M. Krivelevich. On two hamilton cycle problems in random graphs. *Israel Journal of Mathematics*, 166(1):221–234, 2008.
- [9] A. M. Frieze. An algorithm for finding hamilton cycles in random directed graphs. *Journal of Algorithms*, 9(2):181–204, 1988.
- [10] D. Hefetz, A. Steger, and B. Sudakov. Random directed graphs are robustly hamiltonian. *arXiv preprint arXiv:1404.4734*, 2014.
- [11] S. Janson, T. Luczak, and A. Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.
- [12] R. M. Karp. *Reducibility among combinatorial problems*. Springer, 1972.
- [13] M. Krivelevich, C. Lee, and B. Sudakov. Resilient pancyclicity of random and pseudorandom graphs. *SIAM Journal on Discrete Mathematics*, 24(1):1–16, 2010.
- [14] C. Lee and B. Sudakov. Dirac’s theorem for random graphs. *Random Structures & Algorithms*, 41(3):293–305, 2012.
- [15] R. Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv preprint arXiv:1405.6559*, 2014.
- [16] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.
- [17] V. Rödl, A. Ruciński, and E. Szemerédi. An approximate Dirac-type theorem for k -uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.
- [18] B. Sudakov and V. H. Vu. Local resilience of graphs. *Random Structures & Algorithms*, 33(4):409–433, 2008.
- [19] D. B. West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.